

## 機関リポジトリと研究者データベースの連携

森, 雅生  
九州大学大学評価情報室

馬場, 謙介  
九州大学附属図書館研究開発室

<https://doi.org/10.15017/24950>

---

出版情報：九州大学附属図書館研究開発室年報. 2011/2012, pp. 11-13, 2012-07. 九州大学附属図書館  
バージョン：  
権利関係：

## 機関リポジトリと研究者データベースの連携

森 雅生<sup>†</sup> 馬場 謙介<sup>‡</sup>

## &lt;抄録&gt;

本稿は研究者データベースと機関リポジトリの連携を紹介する。研究者データベースに研究業績として登録された学術論文の書誌情報から、機関リポジトリに蓄積されるその論文の本文ファイルへのハイパーリンクの作成を考える。通常、こういったハイパーリンクは、バックグラウンドでの書誌情報の照合により実現される。本稿では、バックグラウンドでの書誌情報の照合を行わない方法を提案している。

<キーワード> 機関リポジトリ, アーカイビング, Web データベース

## Connecting Institutional Repository with Researcher Database

MORI Masao BABA Kensuke

## 1. はじめに

高等教育の内部品質保証は大学の重要な機能のひとつである。教育のみならず、学術研究を行うことも大学の重要な役割である。多くの場合、研究費は政府から配分されるため、研究業績の内容が入手され、それらが開示されることが求められる。この点から、**大学情報**（あるいは、**機関情報**）を幅広く集めることは大学運営にとって必要不可欠である。

一般に、大学はそれぞれ自律的な部局が分散し、協力する組織である。同様に機関情報も大学内に分散している。それらの情報を集めるためには2つの方法がある。ひとつは、既存の分散した情報システムを、統合化されたシステムで新たに置き換えることである。もうひとつは、分散した情報を照合し、紐付けすることである。前者は新しいシステム導入のためのコストとリスクの問題を伴うので、後者の方がより現実的な解法であるといえる。

本稿では、研究者データベースに登録されている学術成果の書誌情報から、機関リポジトリに登録されているその本文データへのハイパーリンクを作成する問題を考える。たいていそういったハイパーリンクの作成は、バックエンドのデータベースでの照合により実現される。しかし、我々は以下の2つの理由から、バックエンドでの書誌情報の照合を行わない手法を提案する。ひとつは、バックエンドでの照合はコストが膨大であるからである。もうひとつは、既に存在するデータだけでなく、これから登録される書誌情報へもリンク機能が自動的に付与されることが求められるからである。

## 2. 問題点

本章では、研究開発の対象である実際のシステムの現状を述べ、我々が取り組む問題点を明らかにする。

## 2.1. 概要

**機関リポジトリ**（以下、**IR**）とは、その機関の学術論文等の研究成果を蓄積し公開するシステムである。IRで学術論文を公開することは、研究成果の社会還元の手段のひとつである。

IRが機能的であるためには、まず、蓄積されている論文の数が十分であるべきである。しかし、ほとんどのIRでは、登録されている論文数は、実際にその機関の研究者が執筆した論文に比べ非常に少ないと推測される[1]。IRの論文数を増やすための第一歩は、これらの「埋もれた」論文の登録を促し、また、新たに執筆された論文が埋もれてしまうのを防ぐことである。このための単純なアプローチは、研究者が論文をIRに登録するための手間を軽減することである。

我々は実際、九州大学のIR (QIR) [2]についてこの問題に直面している。この問題に対する我々の解法は、研究者データベース (DHJS) [3]との連携である。九州大学では、研究成果としての原著論文等の書誌情報の研究者データベースへの登録は義務である。これらの研究者データベースに登録された書誌情報を、IRへの登録に再利用することで、研究者の手間が軽減されることが期待できる。

ほとんどの大学や研究機関は、IRに加えて研究者データベースを保持している。よって、我々のアイデアは他の機関のIRにも適用することができる。

<sup>†</sup> もり まさお 九州大学大学評価情報室 (〒812-8581 福岡市東区箱崎 6-10-1) E-mail: mori@ir.kyushu-u.ac.jp

<sup>‡</sup> ばば けんすけ 九州大学附属図書館研究開発室 (〒812-8581 福岡市東区箱崎 6-10-1) E-mail: baba@lib.kyushu-u.ac.jp

## 2.2. QIR

QIRは九州大学のIRであり、附属図書館によって管理されている。一般に、IRは学術論文等の書誌情報だけでなく、本文ファイルを蓄積する。QIRに登録されている本文ファイルの数は、2012年4月現在で約19,000である。登録論文数は、IRを評価するための重要な指標のひとつである。実際、Ranking Web of World Repositories[5]はそのランキングを決める指標のひとつとして本文ファイルの数を考慮しており、QIRのその数は2012年4月現在で68番目である。そのランキングは約2,000機関のリポジトリを対象にしているため、ほとんどのIRの登録論文数はそれより少ないことになる。

現在、QIRへの論文の登録方法は2つある。

- QIRのアカウントを作成し、登録フォームに沿って書誌情報を入力し、本文ファイルをアップロードする
- 書誌情報と本文ファイルを、QIRの管理者へ電子メールで提出する

QIRの登録フォームでは、タイトルや著者名等の書誌情報を記入しなければならない。普通これらの作業は、研究者が研究成果を自分のwebサイトに公開したり、研究機関に報告書を提出したりする際に繰り返し行われ、これがIRへの登録の障害のひとつであると考えられる。

## 2.3. 共起解析

DHJSは九州大学の研究者データベースであり、大学評価情報室により開発され、管理されている。このデータベースは、原著論文を含む研究者の様々な情報を保持している。DHJSは、入力システムと閲覧システムの2つのサブシステムで構成されている。入力システムは、研究者の研究成果等のDHJSへの登録をサポートしている。閲覧システムは入力システムによってDHJSに登録された情報を表示する。図1はDHJS上に表示された書誌情報のリストの例である。

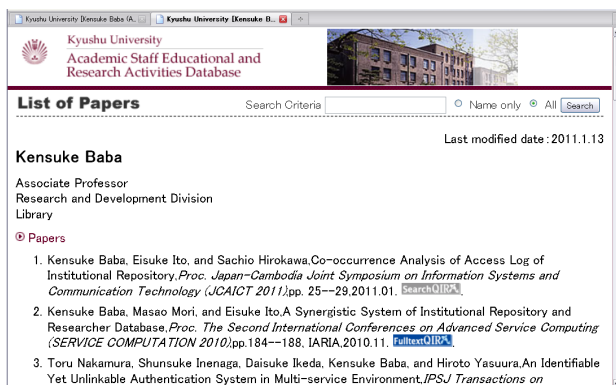


図1: DHJSの論文リストの例

九州大学の研究者は、原著論文の書誌情報を含む研究業績をDHJSに登録しなければならない。DHJSに登録されている書誌情報の数は、2012年4月現在約90,000である。そのうち重複データ（つまり、同一論文の書誌情報）は高々20%と推測される[4]ので、異なる論文の数は約72,000である。一方、前述のように、QIRの登録論文数は約19,000である。つまり、潜在的にはかなりの量の「九州大学で生産されたが、QIRには登録されていない論文」があることになる。

## 3. 解決方法

### 3.1. DHJSからQIRへのリンク

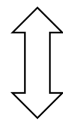
まず、我々はリンクシステムと呼ばれるIR専用の検索エンジンを構築した。論文書誌情報のクエリに対し、リンクシステムはQIR内の関連する論文のURIを返す。

DHJSの論文からの各ハイパーリンクは、書誌情報の項目の集合の形をしている。ハイパーリンクをクリックすることで、利用者はリンクシステムが返す検索結果のURIへ誘導される（図2）。

#### DHJS: Researcher Database

Hyperlink (as a query to the link system)  
<http://linksys.lib.xy-z-u.ac.jp/search.cgi?auth=John+Dow+etal&title=On+a+new+method+of+string+matching&journal=XYZ+pub&date=2011-11>

authors = "John Dow, et. al."  
 title = "On a new method of string matching"  
 journal = "XYZ pub."  
 date = "2011.11"



Possible to create hyperlinks automatically

link forwarded (as the answer from the link system)  
<http://hdl.handle.net/123>

Link system  
 (linksys.lib.xy-z-u.ac.jp)

author\_1 = "John Dow"  
 author\_2 = "Mike Smith"  
 title = "On a new method of string matching"  
 journal = "XYZ pub."  
 year = "2011"  
 month = "November"  
 date = "1"  
 ISBN = "9988776655"  
 URI =  
 "http://hdl.handle.net/123"

#### QIR: Institutional Repository

図2: DHJSからQIRへのリンクシステムの概要

DHJSの書誌情報の中にはQIRのものよりも詳細でないものがある。例えば、共著者による同一論文の書誌情報の重複した登録については、何も処理が施されない。

### 3.2. リンクの動作

我々は、QIRへの論文登録を促すために、QIRとDHJSの閲覧システムを連携させシステムを開発した[6]。このシステムによって、閲覧システムに表示される書誌情報はQIRの対応する本文ファイルへリンクされる。このリンクは閲覧システムに表示される2種類のアイコンで実現される（図1を参照）。各論文について、以下の3つの状態があり得る。

- 青いアイコン「FulltextQIR」が付与されている
- 灰色のアイコン「SearchQIR」が付与されている
- 何も付与されていない

最初の状態はQIRに対応する本文ファイルが存在することを、2番目の状態は本文ファイルが無いことを表している。3番目は著者が本文ファイルへのリンクを付与したくないことを表している。2番目の場合、リンクシステムはQIRでの著者名による検索結果を表示する。加えて、利用者が論文の著者自身の場合、QIRの登録フォームへと誘導する。このとき、DHJSの書誌情報の一部がフォームへの記入のために再利用される。

我々は、このリンクシステムを実際に運用し、アクセスログを解析することで評価を行った。しかし、結果として、論文登録数に関してはこのシステムの効果を見つけることは出来なかった[7]。

#### 4. おわりに

研究者データベースの論文書誌情報から機関リポジトリの本文ファイルへのハイパーリンクを作成する問題に取り組んだ。この問題は、研究者データベースと機関リポジトリとの連携から浮かび上がったものである。

情報開示と大学評価の観点から、研究者データベースは重要でデータが膨大になりつつある。大学が分散した自律的な組織の集まりである以上、大学情報が分散することは大学にとって不可避の問題である。この研究がこの問題の解決につながることを期待する。

#### 参考文献

- [1] K. Baba, N. Hoshiko, E. Kudo, N. Yoshimatsu, and E. Ito. Semi-automated paper-registration system for institutional repository. In Proc. the Third International Conference on Awareness Science and Technology, pages 431–434. IEEE, 2011.
- [2] QIR: 九州大学学術情報リポジトリ. <https://qir.kyushu-u.ac.jp/dspace/>, [Accessed Apr. 30, 2012].
- [3] 九州大学研究者情報. [http://hyoka.ofc.kyushuu-u.ac.jp/search/index\\_e.html](http://hyoka.ofc.kyushuu-u.ac.jp/search/index_e.html), [Accessed Apr. 30, 2012].
- [4] K. Baba, M. Mori, and E. Ito. Identification of scholarly papers and authors by connecting databases. Journal of E-Technology, 2(3):120–125, 2011.
- [5] Ranking Web of World Repositories. <http://repositories.webometrics.info/>, [Accessed Apr. 30, 2012].
- [6] K. Baba, M. Mori, and E. Ito. A synergistic system of institutional repository and researcher database. In Proc. the Second International Conferences on Advanced Service Computing, pages 184–188. IAIRA, 2010.
- [7] K. Baba, T. Tanaka, E. Ishita, M. Mori, E. Ito, and S. Hirokawa. Evaluation of link system between repository and researcher database. In Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation,